

# PATENT ABSTRACTS OF JAPAN

(11) Publication number : 11-331761  
 (43) Date of publication of application : 30. 11. 1999

(51) Int. Cl. H04N 5/91  
 G06F 17/30  
 G06T 13/00

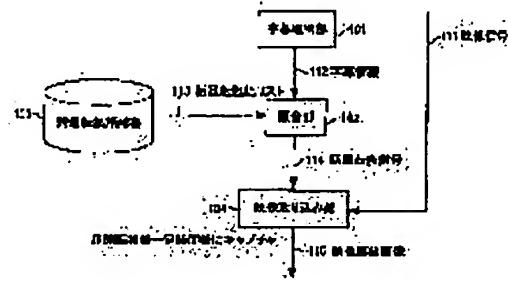
(21) Application number : 10-134125      (71) Applicant : NIPPON TELEGR & TELEPH CORP  
 <NTT>  
 (22) Date of filing : 18. 05. 1998      (72) Inventor : HAYAKAWA KAZUHIRO  
 SUGIZAKI MASAYUKI  
 INAGAKI HIROTO  
 TANAKA KAZUO

(54) METHOD AND DEVICE FOR AUTOMATICALLY SUMMARIZING IMAGE AND RECORDING MEDIUM WITH THE METHOD RECORDED THEREWITH

(57) Abstract:

**PROBLEM TO BE SOLVED:** To provide automatic image summarizing method and device capable of summarizing an image by a less number of images.

**SOLUTION:** A caption acquiring part 101 acquires and outputs caption information 112 synchronized with an image with respect to time. A topic conversion word dictionary 103 outputs a topic conversion word list 113 for switching a topic, such as 'then', 'by the way', 'well', 'continuously', 'next', 'first', and 'so'. A collation part 102 receives the information 112 and the list 113, and when a work in the list 113 is included in the information 112, outputs a topic conversion signal 114. At the receiving of the signal 114, an image entering part 104 enters the current image of a video signal and outputs the entered image as an image which summarizes an image 115. Consequently, the summary of an image can be generated by fewer number of images by collecting the images of each change of a topic of an image with caption information.



LEGAL STATUS

[Date of request for examination]  
 [Date of sending the examiner's decision of rejection]

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平11-331761

(43)公開日 平成11年(1999)11月30日

(51)Int.Cl.  
H 04 N 5/91  
G 06 F 17/30  
G 06 T 13/00

識別記号

F I  
H 04 N 5/91 N  
G 06 F 15/40 3 70 D  
15/401 3 20 A  
15/62 3 40 A

審査請求 未請求 請求項の数9 OL (全9頁)

(21)出願番号 特願平10-134125

(22)出願日 平成10年(1998)5月18日

(71)出願人 000004226

日本電信電話株式会社  
東京都千代田区大手町二丁目3番1号

(72)発明者 早川 和宏  
東京都新宿区西新宿3丁目19番2号 日本  
電信電話株式会社内

(72)発明者 杉崎 正之  
東京都新宿区西新宿3丁目19番2号 日本  
電信電話株式会社内

(72)発明者 稲垣 博人  
東京都新宿区西新宿3丁目19番2号 日本  
電信電話株式会社内

(74)代理人 弁理士 志賀 富士介 (外1名)

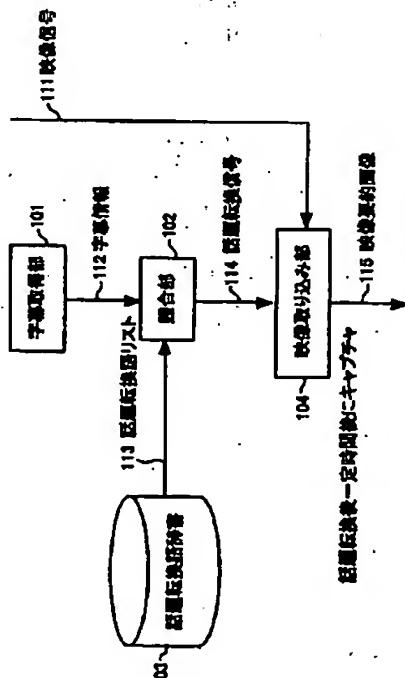
最終頁に続く

(54)【発明の名称】 映像の自動要約方法およびその装置およびその方法を記録した記録媒体

(57)【要約】

【課題】 より少ない数の画像で映像を要約することができる映像の自動要約方法およびその装置を提供する。

【解決手段】 字幕取得部101は、映像と時間的に同期した字幕情報112を取得して出力する。話題転換語辞書103は、「次に」「ところで」「さて」「統いでは」「次は」「まず」「それでは」等、話題を切り替え用の語話題転換語リスト113を出力する。照合部102は、字幕情報112と話題転換語リスト113を受け取り、話題転換語リスト113中の語が字幕情報112に含まれていれば、話題転換信号114を出力する。映像取り込み部104は、話題転換信号114を受け取ると映像信号111のその時点での画像を取り込み、映像要約画像115として出力する。これにより、字幕情報付き映像の話題の変わり目毎の画像を集めることで、より少ない数の画像で映像の要約を生成する。



## 【特許請求の範囲】

【請求項1】 映像の内容を自動的に要約するための方  
法において、  
映像と同期する字幕情報を取得し、  
前記取得された字幕情報中の話題転換語の有無を検出  
し、  
前記話題転換語が検出されたことを条件として映像信号  
から画像を取り込み出力する、  
ことを特徴とする映像の自動要約方法。

【請求項2】 映像の内容を自動的に要約するための方  
法において、  
映像と同期する字幕情報を取得し、  
前記取得された字幕情報のテキストから形態素解析によ  
り単語を抽出し、  
前記抽出された単語の中での話題転換語の有無を検出  
し、  
前記話題転換語が検出されたことを条件として映像信号  
から画像を取り込み出力する、  
ことを特徴とする映像の自動要約方法。

【請求項3】 前記話題転換語が検出されたことを条件  
として映像信号から画像を取り込み出力する過程では、  
話題転換語が検出されたら映像信号のショット切り替え  
検出を開始し、  
前記ショット切り替えが検出されたら一定時間後の映像  
信号から画像を取り込み出力する、  
ことを特徴とする請求項1または請求項2に記載の映像  
の自動要約方法。

【請求項4】 映像の内容を自動的に要約するための装  
置において、  
映像と同期する字幕情報を取得する手段と、  
前記取得された字幕情報中の話題転換語の有無を検出する  
手段と、  
前記話題転換語が検出されたことを条件として映像信号  
から画像を取り込み出力する手段と、  
を具備することを特徴とする映像の自動要約装置。

【請求項5】 映像の内容を自動的に要約するための装  
置において、  
映像と同期する字幕情報を取得する手段と、  
前記取得された字幕情報のテキストから形態素解析によ  
り単語を抽出する手段と、  
前記抽出された単語の中での話題転換語の有無を検出する  
手段と、  
前記話題転換語が検出されたことを条件として映像信号  
から画像を取り込み出力する手段と、  
を具備することを特徴とする映像の自動要約装置。

【請求項6】 前記話題転換語が検出されたことを条件  
として映像信号から画像を取り込み出力する手段は、  
話題転換語が検出されたら映像信号のショット切り替え  
検出を開始する手段と、  
ショット切り替えが検出されたら一定時間後の映像信号

から画像を取り込み出力する手段と、  
を具備することを特徴とする請求項4または請求項5に  
記載の映像の自動要約装置。

【請求項7】 映像の内容を自動的に要約するための方  
法における、  
映像と同期する字幕情報を取得する手順と、  
前記取得された字幕情報中の話題転換語の有無を検出する  
手順と、  
前記話題転換語が検出されたことを条件として映像信号  
から画像を取り込み出力する手順と、  
をコンピュータに実行させるためのプログラムを、該コ  
ンピュータが読み取り可能な記録媒体に記録した、  
ことを特徴とする映像の自動要約方法を記録した記録媒  
体。

【請求項8】 映像の内容を自動的に要約するための方  
法における、  
映像と同期する字幕情報を取得する手順と、  
前記取得された字幕情報のテキストから形態素解析によ  
り単語を抽出する手順と、  
前記抽出された単語の中での話題転換語の有無を検出する  
手順と、  
前記話題転換語が検出されたことを条件として映像信号  
から画像を取り込み出力する手順と、  
をコンピュータに実行させるためのプログラムを、該コ  
ンピュータが読み取り可能な記録媒体に記録した、  
ことを特徴とする映像の自動要約方法を記録した記録媒  
体。

【請求項9】 前記話題転換語が検出されたことを条件  
として映像信号から画像を取り込み出力する手順では、  
話題転換語が検出されたら映像信号のショット切り替え  
検出を開始し、  
前記ショット切り替えが検出されたら一定時間後の映像  
信号から画像を取り込み出力する、  
ことを特徴とする請求項7または請求項8に記載の映像  
の自動要約方法を記録した記録媒体。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】本発明は、映像中の一部の画  
像を取り出すことで映像の要約を生成する映像の自動要  
約方法およびその装置に関する。

## 【0002】

【従来の技術】従来、映像を要約するために映像中の画  
像を選択的に取り出すための方法としては、ショットの  
切り替わりを検出して各ショットごとに画像を一つ取り  
出す方法があった。

## 【0003】

【発明が解決しようとする課題】しかしながら、通常、  
上記従来の方法では数秒に一つの画像が生成され、映像  
を要約するには画像の数が多くなりすぎるという問題が  
あった。

【0004】そこで本発明の課題は、より少ない数の画

3

像で映像を要約することができる映像の自動要約方法およびその装置を提供することにある。

【0005】

【課題を解決するための手段】本発明は、以下に列記する発明により上記の課題を解決する。

【0006】本発明(1)は、映像と同期する字幕情報を取得し、前記取得された字幕情報中の話題転換語の有無を検出し、前記話題転換語が検出されたら映像信号から画像を取り込み出力する映像の自動要約方法である。

【0007】本発明(2)は、映像と同期する字幕情報を取得し、前記取得された字幕情報のテキストから形態素解析により単語を抽出し、前記抽出された単語の中での話題転換語の有無を検出し、前記話題転換語が検出されたら映像信号から画像を取り込み出力する映像の自動要約方法である。

【0008】本発明(3)は、上記の発明(1)または(2)において、話題転換語が検出されたら、すぐには映像信号から画像を取り込みます、映像信号のショット切り替え検出を開始し、ショット切り替えが検出されたら一定時間後の映像信号から画像を取り込み出力する映像の自動要約方法である。

【0009】本発明(4)は、映像と同期する字幕情報を取得する手段と、前記取得された字幕情報中の話題転換語の有無を検出する手段と、前記話題転換語が検出されたら映像信号から画像を取り込み出力する手段と、を具備する映像の自動要約装置である。

【0010】本発明(5)は、映像と同期する字幕情報を取得する手段と、前記取得された字幕情報のテキストから形態素解析により単語を抽出する手段と、前記抽出された単語の中での話題転換語の有無を検出する手段と、前記話題転換語が検出されたら映像信号から画像を取り込み出力する手段と、を具備する映像の自動要約装置である。

【0011】本発明(6)は、上記の発明(4)または(5)において、前記話題転換語が検出されたら映像信号から画像を取り込み出力する手段が、話題転換語が検出されたら映像信号のショット切り替え検出を開始する手段と、ショット切り替えが検出されたら一定時間後の映像信号から画像を取り込み出力する手段と、を具備する映像の自動要約装置である。

【0012】本発明(7)は、映像の内容を自動的に要約するための方法における、映像と同期する字幕情報を取得する手順と、前記取得された字幕情報中の話題転換語の有無を検出する手順と、前記話題転換語が検出されたことを条件として映像信号から画像を取り込み出力する手順と、をコンピュータに実行させるためのプログラムを、該コンピュータが読み取り可能な記録媒体に記録した映像の自動要約方法を記録した記録媒体である。

【0013】本発明(8)は、映像の内容を自動的に要約するための方法における、映像と同期する字幕情報を

取得する手順と、前記取得された字幕情報のテキストから形態素解析により単語を抽出する手順と、前記抽出された単語の中での話題転換語の有無を検出する手順と、前記話題転換語が検出されたことを条件として映像信号から画像を取り込み出力する手順と、をコンピュータに実行させるためのプログラムを、該コンピュータが読み取り可能な記録媒体に記録した映像の自動要約方法を記録した記録媒体である。

【0014】本発明(9)は、上記の発明(7)、(8)において、上記話題転換語が検出されたことを条件として映像信号から画像を取り込み出力する手順では、話題転換語が検出されたら映像信号のショット切り替え検出を開始し、前記ショット切り替えが検出されたら一定時間後の映像信号から画像を取り込み出力する自動要約方法を記録した記録媒体である。

【0015】本発明(1)、(4)、(7)は、字幕情報中に話題転換語が文字列として検出される都度、その瞬間の映像を静止画として出力することを最も主要な特徴とする。本発明(1)、(4)、(7)によれば、字幕情報付き映像の話題の変わり目ごとの画像を集めることで、より少ない数の画像で映像の要約を生成することができる。

【0016】本発明(2)、(5)、(8)は、字幕情報中に話題転換語が単語として検出される都度、その瞬間の映像を静止画として出力することを最も主要な特徴とする。本発明(2)、(5)、(8)によれば、字幕情報付き映像の話題の変わり目を精度良く検出して、話題の変わり目ごとの画像を集めることができ、より少ない数の画像で映像の要約を生成することができる。

【0017】本発明(3)、(6)、(9)は、字幕情報中もしくはその中の単語中に話題転換語が検出される都度、その直後に開始された映像ショット中の画像を静止画として出力することを最も主要な特徴とする。本発明(3)、(6)、(9)によれば、字幕情報付き映像において話題が変わった後の最初の映像ショットの画像を集めることで、より少ない数の、よりふさわしい画像で映像の要約を生成することができる。

【0018】

【発明の実施の形態】以下、本発明の実施の形態について図を用いて詳細に説明する。

【0019】図1は、本発明(4)の一実施形態例を示す装置の構成を表すブロック図である。また、図2は、本発明(1)の一実施形態例の方法を示すとともに、上記装置の動作を表すフロー図である。

【0020】本実施形態例において、字幕取得部101は、映像と時間的に同期した字幕情報112を取得して出力する。話題転換語辞書103は、話題転換語リスト113を出力する。照合部102は、字幕情報112と話題転換語リスト113を受け取り、話題転換語リスト113中の語が字幕情報112に含まれていれば、話題

転換信号114を出力する。映像取り込み部104は、話題転換信号114を受け取ると映像信号111のその時点での画像を取り込み、映像要約画像115として出力する。以上の各部での処理を終了まで繰り返すことで、映像の要約を得る。

【0021】ここで話題転換語辞書103が出力する話題転換語リスト113とは、「次に」「ところで」「さて」「続いては」「次は」「ます」「それでは」など、話題を切り替えるために用いられる語のリストである。

【0022】字幕取得部101における処理内容は字幕の提供形態によるが、一例を挙げればテレビ放送で行われている字幕はビデオ信号中の垂直ランキング区間の走査線内に格納されている。したがって、この場合は字幕取得部101は映像信号111から字幕を取得することになる。

【0023】図3は図1の装置の動作および図2の方法の作用を説明するための図である。この図は、テレビのニュース番組とその字幕情報からどのように映像要約画像を出力するかを表している。

【0024】図3(a)の3つの画面は上から下へ字幕および映像内容の移り変わりを示している。最初の画面の字幕は、話題転換語が含まれていないので何もしない。字幕が切り替わって2番目の画面になると、「次に」が話題転換語に一致するので、話題転換信号が出力される。話題転換信号により映像取り込み部でその瞬間の画像(図3(a)の2番目の画像)が取り込まれ、出力される。さらに字幕が切り替わって3番目の画面になると、ここでは話題転換語は存在しないので何もしない。

【0025】取り込まれた画像は、例えば図3(b)の出力例のように並べてみせることで、映像の内容を大まかに示すことができる。

【0026】図4は、本発明(5)の一実施形態例を示す装置の構成を表すブロック図である。また、図5は、本発明(2)の一実施形態例の方法を示すとともに、上記装置の動作を表すフロー図である。

【0027】本実施形態例での、構成と動作は、ほぼ図1の装置、図2の方法と同様であるが、形態素解析部301により字幕中に含まれる単語リストを抽出することで、話題転換の検出精度を向上させる。

【0028】図1の装置、図2の方法では、たとえば「まずいことに」という字幕を「まず」という話題転換語と照合した場合、「まず」という文字列が字幕中に含まれているために、これを誤って話題転換語として検出してしまう問題がある。

【0029】これに対し本装置、本方法では、形態素解析部301で「まずいことに」を「まずい」「こと」「に」という単語に分解する。照合部102は、単語レベルで照合を行う。そのため「まずい」は「まず」とは一致せず、話題転換語辞書113に含まれていない言葉

を話題転換語として検出してしまってはならない。

【0030】図6は、本発明(6)の一実施形態例を示す装置の構成を表すブロック図である。また、図7は、本発明(3)の一実施形態例の方法を示すとともに、上記装置の動作を表すフロー図である。

【0031】本実施形態例において、字幕取得部401は映像と時間的に同期した字幕情報411を取得して出力する。話題転換語辞書404は話題転換語リスト413を出力する。照合部403は字幕情報411と話題転換語リスト413を受け取り、話題転換語リスト413中の語が字幕情報411に含まれていれば、話題転換信号414を出力してショット切替検出部405を起動するとともに、並行して字幕取得部401からの話題転換の検出の処理を終了まで繰り返す(図7(a))。ショット切替検出部405は、話題転換信号414を受け取ると映像信号415を監視し、最初にショット切替が起ったときにショット切替信号416を出力する。映像取り込み部406はショット切替信号416を受け取ると、それから一定時間後の映像信号415中の画像を取り込み、映像要約画像417として出力して、ショット切替検出と映像信号からの画像の取り込みの処理を終了する。ショット切替検出部405がショット切替を検出しない場合には、次の話題転換が検出されたとき、および一連の話題転換の検出の処理が終了したとき、終了する(図7(b))。

【0032】本装置の特徴は、話題転換検出後すぐの画像ではなく、次に映像が切り替わった後の画像を用いる点である。ここで、話題転換検出後すぐに画像を取り込まないのは、話題転換と新しい話題に関する映像との間に時間的にずれがある可能性があるからである。たとえば図3(a)はニュース番組であるが、まずアナウンサーが話題転換語を用いて新しい話題を導入し、次にショットが切り替わってその話題に関する実際の映像が流れている構成になっている。映像内容の要約としてよりふさわしいのは、2番目のアナウンサーの画像ではなく、3番目の話題に関する映像のほうであると考えられる。

【0033】また、ショット切替は、画像の特殊効果(ワイプ、ディゾルブなど)を伴うことがあり、ショット切替の検出直後の画像は乱れている可能性がある。そのため、ショット切替検出から実際の画像取り込みまでにタイムラグを設けている。このタイムラグの大きさとしては、1ないし数秒が適当である。

【0034】なお、本発明(3)、(6)の実施形態例で示したショット切替後の画像を要約を生成するために取り込むことは、本発明(1)、(4)にも適用することが可能である。

【0035】また、図1、図4、図6で示した装置の各部の一部もしくは全部を、コンピュータを用いて機能させることができること、あるいは、図2、図5、図7で

示した処理の手順をコンピュータで実行させることができることは言うまでもなく、コンピュータをその装置の各部として機能させるためのプログラム、あるいは、コンピュータでその処理の手順を実行させるためのプログラムを、そのコンピュータが読み取り可能な記録媒体、例えば、FD(フロッピーディスク)や、MO、ROM、メモリカード、CD、DVD、リムーバブルディスクなどに記録して提供し、配布することが可能である。

#### 【 0036 】

【 発明の効果】以上説明したように、本発明(1)、(4)、(7)を用いれば、字幕情報付き映像の話題の変わり目ごとの画像を集めることで、より少ない数の画像で映像の要約を生成することができる。

【 0037 】本発明(2)、(5)、(8)を用いれば、字幕情報付き映像の話題の変わり目を精度良く検出して、話題の変わり目ごとの画像を集めることができ、より少ない数の画像で映像の要約を生成することができる。

【 0038 】本発明(3)、(6)、(9)を用いれば、字幕情報付き映像において話題が変わった後の最初の映像ショットの画像を集めることで、より少ない数の、よりふさわしい画像で映像の要約を生成することができる。

#### 【 図面の簡単な説明】

【 図1】本発明(4)の一実施形態例を示す装置の構成を表すブロック図である。

【 図2】本発明(1)の一実施形態例の方法を示すとともに、図1の装置の動作を表すフロー図である。

【 図3】(a)、(b)は、図1の装置の動作、図2の方法の作用を説明するための図である。

【 図4】本発明(5)の一実施形態例を示す装置の構成を表すブロック図である。

【 図5】本発明(2)の一実施形態例の方法を示すとともに、図4の装置の動作を表すフロー図である。

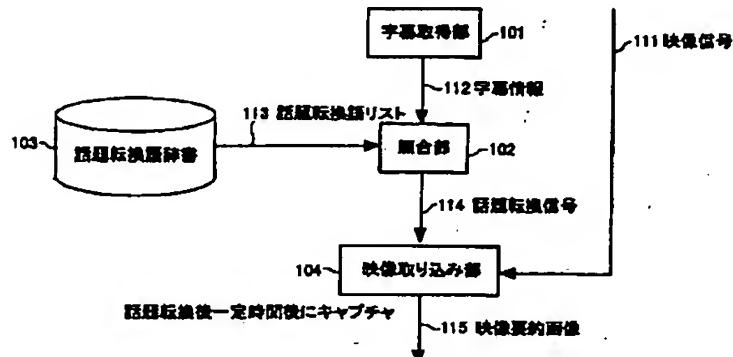
【 図6】本発明(6)の一実施形態例を示す装置の構成を表すブロック図である。

【 図7】(a)、(b)は、本発明(3)の一実施形態例の方法を示すとともに、図6の装置の動作を表すフロー図である。

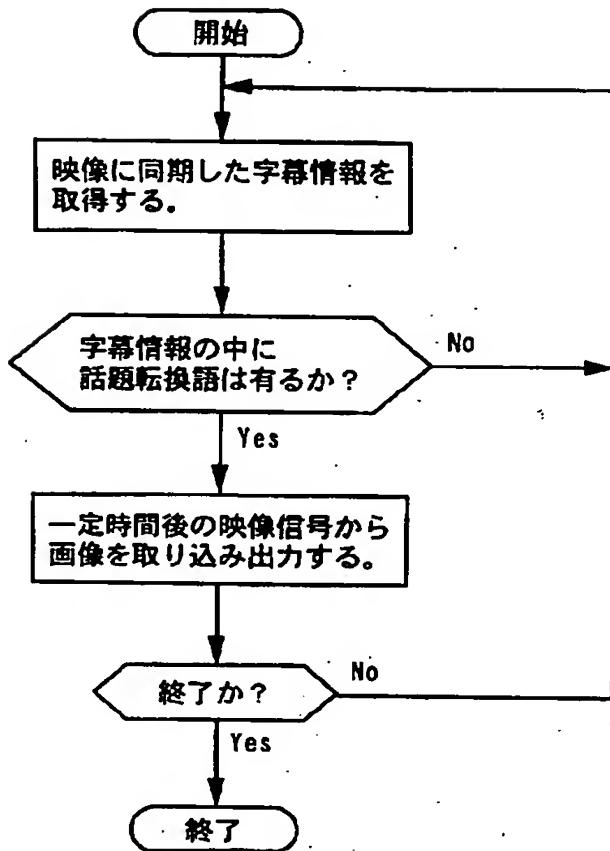
#### 【 符号の説明】

- |     |            |
|-----|------------|
| 101 | …字幕取得部     |
| 102 | …照合部       |
| 103 | …話題転換語辞書   |
| 104 | …映像取り込み部   |
| 111 | …映像信号      |
| 112 | …字幕情報      |
| 113 | …話題転換語リスト  |
| 114 | …話題転換信号    |
| 301 | …形態素解析部    |
| 311 | …字幕中単語リスト  |
| 401 | …字幕取得部     |
| 402 | …形態素解析部    |
| 403 | …照合部       |
| 404 | …話題転換後辞書   |
| 405 | …ショット切替検出部 |
| 406 | …映像取り込み部   |
| 411 | …字幕情報      |
| 412 | …字幕中単語リスト  |
| 413 | …話題転換語リスト  |
| 414 | …話題転換信号    |
| 415 | …映像信号      |
| 416 | …ショット切替信号  |
| 417 | …映像要約画像    |

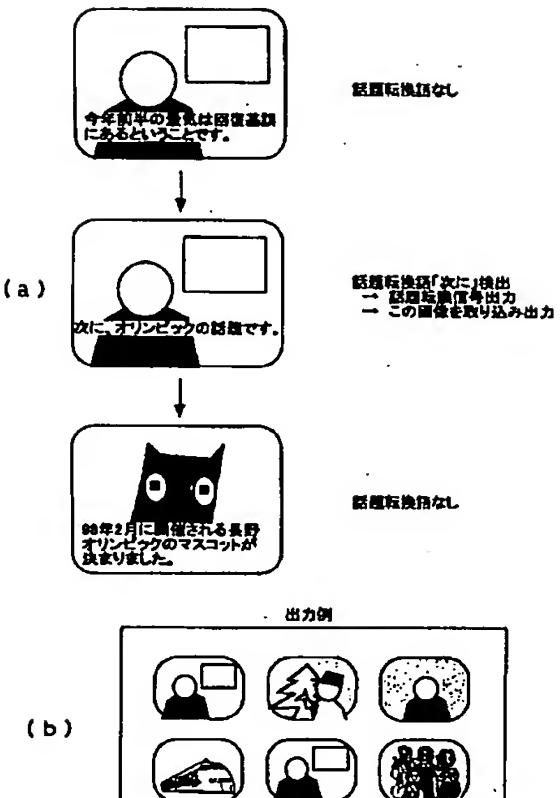
【 図1】



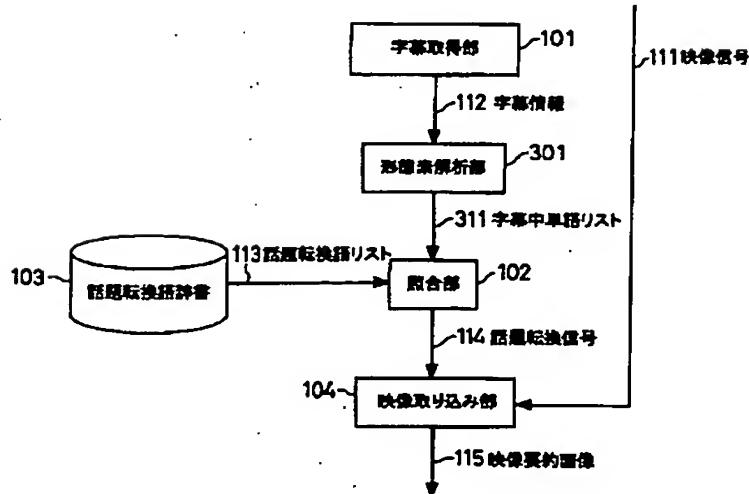
【 図2 】



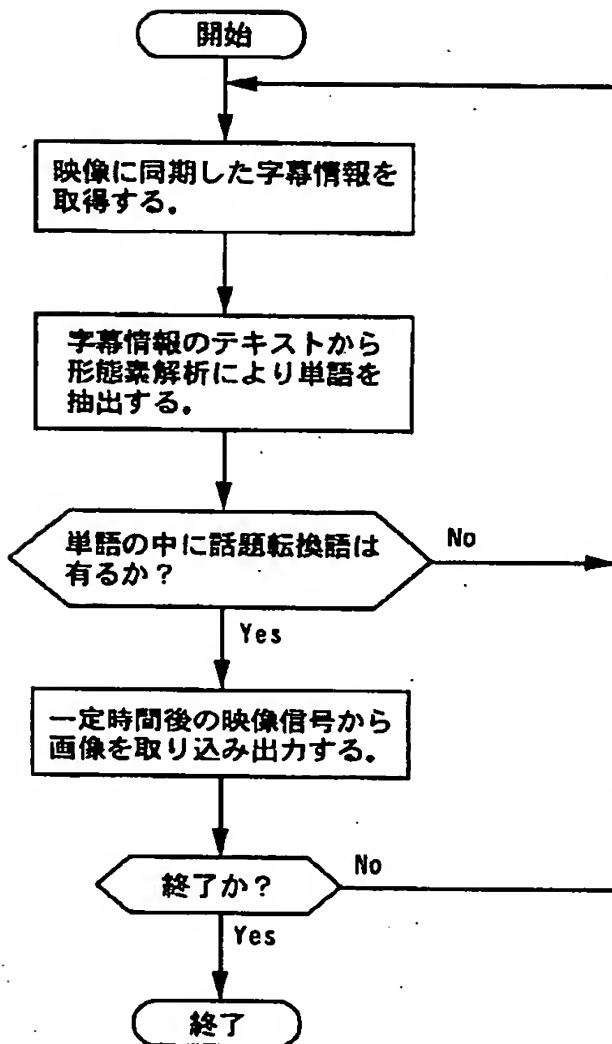
【 図3 】



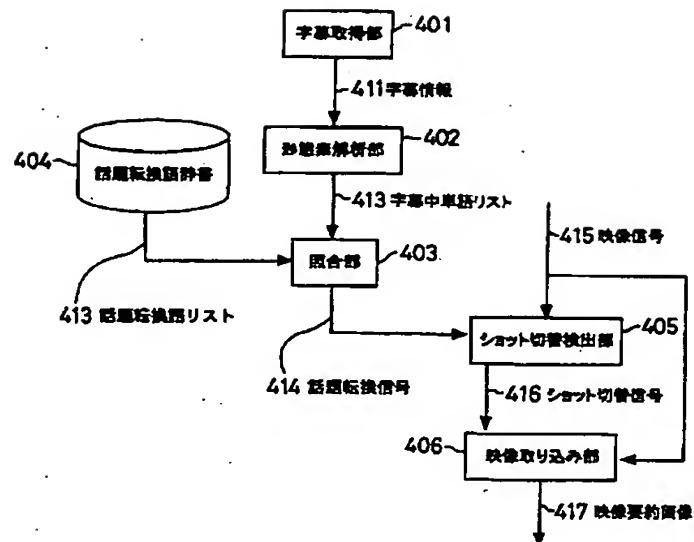
【 図4 】



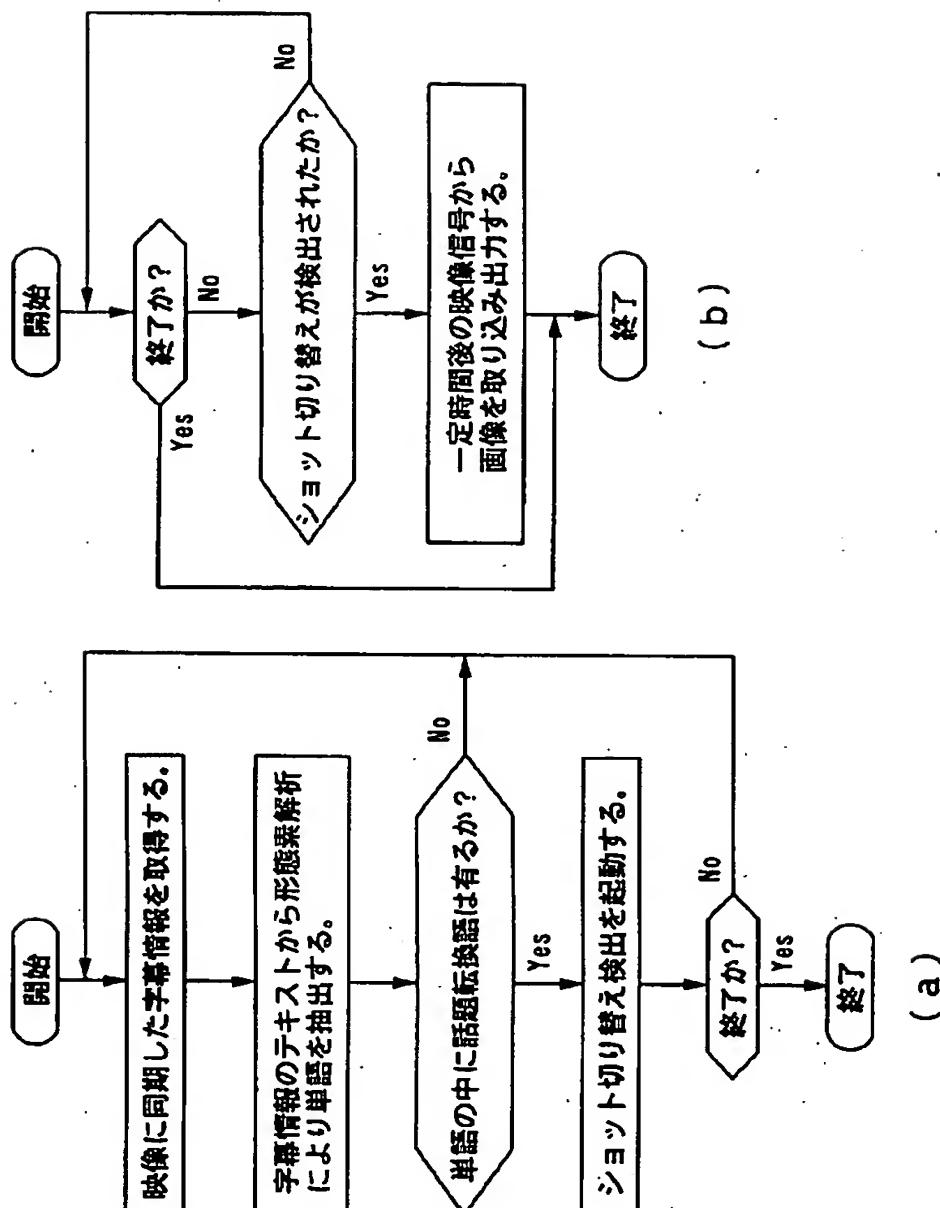
【 図5 】



【 図6 】



【 図7 】



フロントページの続き

(72)発明者 田中 一男

東京都新宿区西新宿3丁目19番2号 日本  
電信電話株式会社内